

# Supplementary Materials for *PASSION: Towards Effective Incomplete Multi-Modal Medical Image Segmentation with Imbalanced Missing Rates*

Anonymous Authors

## 1 COMPLETE QUALITATIVE COMPARISON

Exemplar segmentation results on BraTS2020 and MyoPS2020 under certain modality combinations are presented in Sections 4.3 and 4.4. For a more comprehensive evaluation, qualitative segmentation results of all modality combinations are illustrated in Figures 1 and 2. Compared to the baseline and other comparison approaches, PASSION consistently achieves performance improvements, leading to the fewest false positives and false negatives. It shows PASSION’s superiority in dealing with various modality-missing scenarios.

## 2 ABLATION STUDY ON $D_n^m$

In Section 3.3, we use the distance  $D_n^m$  defined as L2-Proto-Distance across all classes to represent the relative preference of the multi-modal teacher to each uni-modal  $m$ . Here, we adopt various calculations of  $D_n^m$  for comparison to verify the validity of L2-Proto-Distance, including the uni-modal’s Dice loss, KL loss, and Proto loss as the measure of  $D_n^m$  respectively. Quantitative comparison results are summarized in Table 1. In general, compared to the baseline, regularization through re-balancing relative preference is beneficial, proving the effectiveness of PASSION. Compared to other calculations, L2-Proto-Distance is a better measure of relative reference, achieving the best overall segmentation performance.

**Table 1: Ablation study on  $D_n^m$  in Eq. 7 evaluated on BraTS2020 under  $MR = (0.2, 0.4, 0.6, 0.8)$ .**

Distance	DSC [%] ↑				HD [mm] ↓			
	WT	TC	ET	Avg.	WT	TC	ET	Avg.
None	83.39	70.37	52.15	68.64	12.59	13.85	9.25	11.90
Dice Loss	83.37	70.53	52.12	68.67	13.22	12.42	8.72	11.45
KL Loss	83.56	70.58	51.89	68.68	12.97	12.21	8.85	11.34
Proto Loss	83.55	<b>71.31</b>	52.44	69.10	13.69	12.78	<b>8.14</b>	11.54
L2-Proto-Distance	<b>83.91</b>	71.15	<b>52.77</b>	<b>69.28</b>	<b>11.92</b>	<b>11.78</b>	8.42	<b>10.71</b>

## 3 ABLATION STUDY ON $\tau$

The temperature  $\tau$  in Eq. 3 is a hyper-parameter controlling the softness of probability distributions. To study the effect of  $\tau$  on segmentation, ablation studies are conducted by varying the value of  $\tau$ . In addition, to validate the proposed Pixel-wise Self-Distillation, we further change the KL loss into a weighted cross-entropy (WCE) loss with ground-truth and a Feature-wise L2-loss (F-L2) as a comparison. Quantitative results are summarized in Table 2. In general, penalizing pixel-wise self-distillation through the KL loss is more beneficial and stable, being relatively insensitive to  $\tau$ . Specifically, setting  $\tau = 4$  and  $\tau = 3$  achieves the best Dice and HD performance.

## 4 ADDITIONAL IMPLEMENTATION DETAILS

In Section 3.3, we mentioned the gradient-wise re-balancing regularization by applying the gradient descent of epoch-average relative

**Table 2: Ablation study on  $\tau$  in Eq. 3.**

$\tau$	DSC [%] ↑				HD [mm] ↓			
	WT	TC	ET	Avg.	WT	TC	ET	Avg.
1	83.01	69.12	51.82	67.98	14.50	13.03	8.73	12.09
2	83.51	70.63	51.91	68.68	12.22	12.17	8.07	10.82
3	83.66	70.99	52.46	69.04	11.80	<b>10.87</b>	8.01	<b>10.23</b>
4	<b>83.91</b>	<b>71.15</b>	<b>52.77</b>	<b>69.28</b>	12.59	13.85	8.42	11.62
5	82.73	69.80	52.43	68.32	14.06	12.95	9.00	12.00
6	83.13	69.86	52.47	68.49	12.96	12.32	7.90	11.06
7	82.87	70.18	51.81	68.29	13.48	12.24	8.04	11.25
8	83.13	70.31	52.00	68.48	12.06	11.57	7.92	10.52
9	83.23	70.51	52.53	68.76	<b>11.64</b>	11.70	<b>7.51</b>	10.28
10	83.44	70.14	52.25	68.61	12.90	11.95	7.82	10.89
WCE	80.83	70.04	52.04	67.64	20.13	15.40	9.54	15.02
F-L2	81.91	67.72	50.25	66.63	14.48	13.09	8.47	12.01

preference. As the gradient does not directly correspond to the Pixel-wise Self-Distillation optimization objective, for further constraints, we used projected gradient descent instead, formulated as:

$$\begin{aligned} \beta_{r+1}^m &= \beta_r^m - \gamma \cdot \bar{R}^p, \\ [\beta_{r+1}^1, \beta_{r+1}^2, \dots, \beta_{r+1}^M] &= Proj_{\Gamma_\beta} [\beta_{r+1}^1, \beta_{r+1}^2, \dots, \beta_{r+1}^M], \end{aligned} \quad (1)$$

where  $r$  indexes the epoch,  $\Gamma_\beta$  represents the feasible domain of  $\beta = [\beta^1, \beta^2, \dots, \beta^M]$ . Here we set  $\Gamma_\beta = \{\beta \in \mathbb{R}^M | \beta^m \geq 0.1, \|\beta\|_2 = \|\beta_0\|_2\}$ .  $Proj_{\Gamma_\beta}[\cdot]$  represents the projection to domain  $\Gamma_\beta$ .  $\gamma$  is the learning rate of the weighted coefficients.

In Section 4.6, we conducted a plug-and-play ablation study under various SOTA networks. To ensure the fairness of comparison experiments, we modified different networks into a uniform size (i.e., the basic dimension size is set as 8), focusing more on the effectiveness of PASSION in modality-fusion methods. Different from classification tasks, in all segmentation experiments, we set the batch size as 1 (following previous works based on 1 or 2). Such a setup makes models friendly for real-missing input. A larger batch size can be realized by gradient accumulation, which may improve the overall performance but is out of our focus. In future studies, how to make use of the batch samples under *IDT* with imbalanced missing rates may deserve further exploration.

## 5 HD PERFORMANCE ON BRATS2020

In addition to the Dice performance evaluation on BraTS2020 in Section 4.6, the HD performance comparison is summarized in Table 3. Consistent with the observations in Section 4.6, *IDT* brings severe performance degradation compared to *PDT*. When extending PASSION to different backbones, consistent performance improvements are achieved, demonstrating its effectiveness in shape preservation.

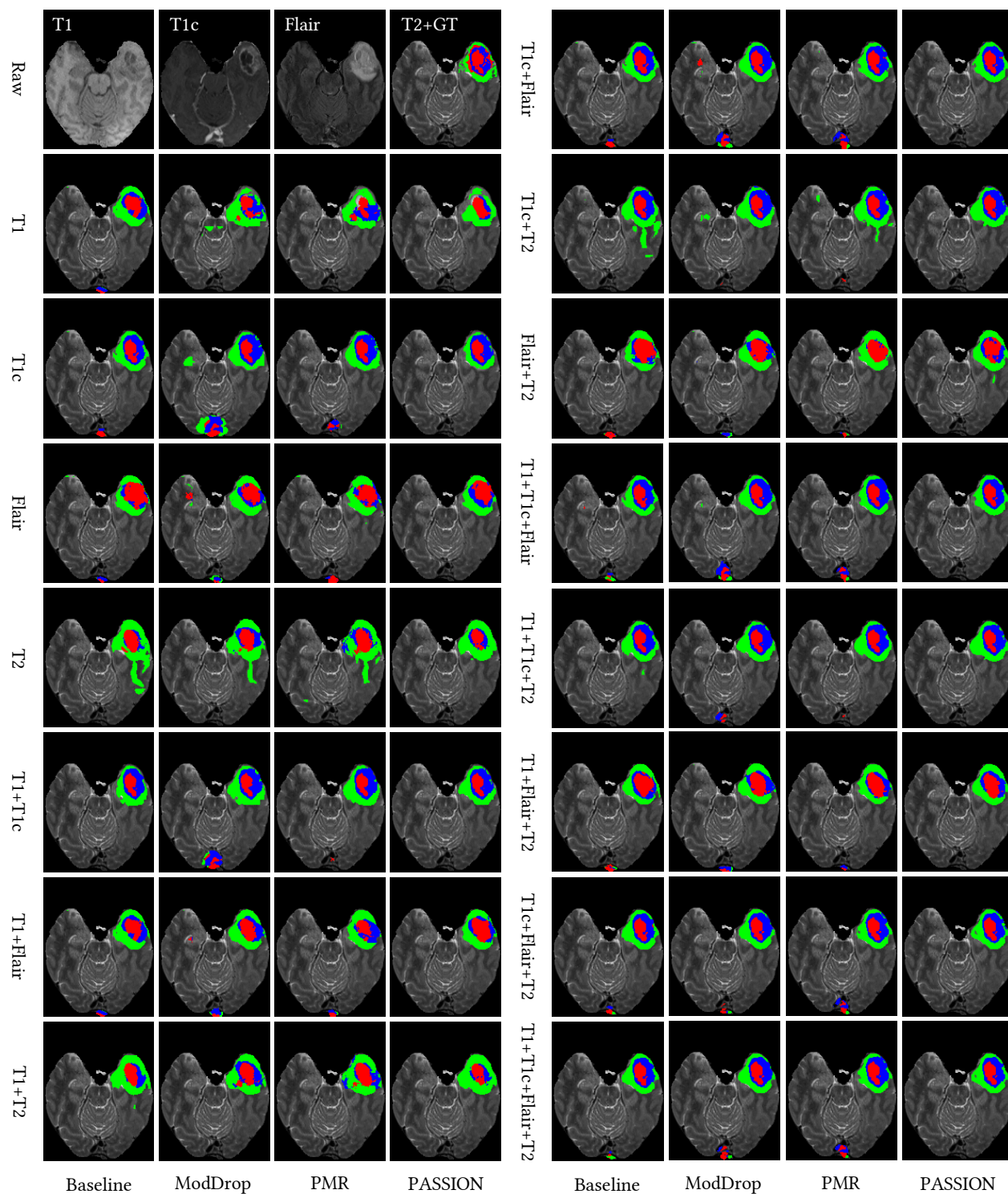


Figure 1: Complete qualitative comparison on BraTS2020.

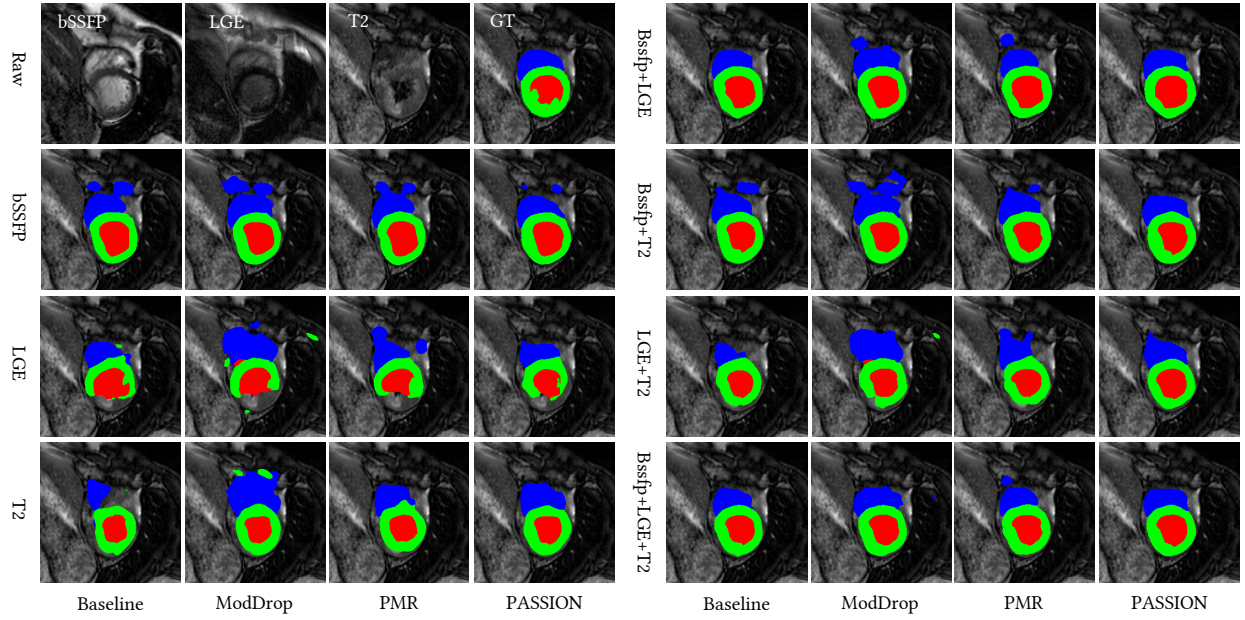


Figure 2: Complete qualitative comparison on MyoPS2020.

Table 3: Quantitative comparison (measured by HD (mm)) on various backbones evaluated on BraTS2020 under the *IDT* (i.e., imperfect data training) setting  $MR = (0.2, 0.4, 0.6, 0.8)$  and the *PDT* (i.e., perfect data training) setting  $MR = (0, 0, 0, 0)$  for T1, T1c, Flair, and T2 respectively.

Type	Setting	T1 T1c Flair T2	● ○ ○ ○	○ ● ○ ○	○ ○ ● ○	○ ○ ○ ●	● ● ○ ○	● ○ ● ○	● ○ ○ ○	○ ● ○ ○	○ ● ○ ○	○ ○ ○ ●	● ● ○ ○	● ● ○ ○	○ ○ ○ ●	○ ○ ○ ●	○ ○ ○ ●	Avg.
WT	PDT	RFNet	33.15	36.06	27.73	23.68	26.99	22.42	16.69	23.31	20.76	18.89	22.24	18.14	17.99	20.91	17.95	23.13
		mmFormer	27.45	28.12	17.82	18.71	21.15	12.83	12.65	11.56	11.03	12.93	10.36	9.33	10.65	10.88	9.37	14.99
		M2FTrans	30.45	36.32	31.67	26.73	25.67	20.85	20.85	26.34	24.95	18.68	20.71	19.93	15.47	19.74	20.04	23.89
	IDT	RFNet	28.61	39.54	36.93	43.12	26.98	23.19	28.38	29.10	34.76	28.83	24.56	28.54	20.69	28.61	26.79	29.91
		mmFormer	27.65	30.79	35.13	40.32	17.89	24.11	21.95	29.68	38.23	26.74	21.19	23.32	20.05	28.03	20.81	27.06
		M2FTrans	81.11	35.06	33.74	42.08	15.97	19.69	20.14	28.56	35.37	29.16	21.22	21.30	19.44	27.41	20.71	30.06
	IDT (+PASSION)	RFNet	16.31	17.33	19.95	16.88	24.63	16.26	16.37	19.50	22.57	16.80	17.56	23.75	14.48	18.96	17.09	18.56
		mmFormer	14.86	15.05	17.26	11.79	13.15	11.02	9.88	16.15	14.35	10.40	10.79	11.17	9.53	12.42	11.07	12.59
		M2FTrans	14.30	15.65	15.53	12.25	11.69	12.93	10.46	18.03	12.34	16.36	15.61	12.13	12.52	19.03	17.08	14.39
TC	PDT	RFNet	24.95	33.92	29.80	21.37	22.09	20.00	17.37	22.92	20.15	20.82	21.05	16.59	19.40	19.57	18.20	21.88
		mmFormer	26.08	30.67	25.65	20.10	15.96	18.17	15.14	15.00	11.63	16.37	13.41	10.81	13.83	13.11	11.80	17.18
		M2FTrans	28.46	32.09	29.40	23.68	15.98	18.79	16.97	25.75	24.06	19.44	15.66	18.04	15.55	25.27	19.21	21.89
	IDT	RFNet	20.61	42.65	42.53	30.33	27.81	24.11	28.62	29.94	19.22	29.08	20.72	23.92	23.55	26.07	26.48	27.71
		mmFormer	23.45	24.08	40.19	32.98	8.06	27.60	16.64	26.79	18.08	29.43	21.24	8.60	22.58	20.06	14.69	22.30
		M2FTrans	93.35	36.93	34.90	28.04	12.14	25.50	14.56	27.93	18.78	25.96	21.05	12.42	22.77	24.37	17.38	27.74
	IDT (+PASSION)	RFNet	15.32	16.30	17.08	14.54	17.55	16.10	15.36	14.94	16.83	15.33	14.48	15.12	14.32	13.28	10.51	15.14
		mmFormer	14.29	14.57	19.13	12.92	8.71	13.72	12.86	22.85	9.77	15.94	15.44	8.42	13.30	13.06	12.72	13.85
		M2FTrans	12.40	12.57	14.62	13.05	7.71	13.70	10.67	13.37	9.04	14.67	13.01	8.06	12.44	11.78	10.49	11.84
ET	PDT	RFNet	19.00	23.07	21.86	14.48	11.20	14.02	11.42	12.41	9.57	13.81	10.66	6.82	11.54	12.21	9.62	13.45
		mmFormer	20.31	15.83	18.40	13.32	9.60	14.82	10.57	10.45	8.47	12.58	9.44	8.01	11.21	7.78	7.43	11.88
		M2FTrans	18.59	19.79	18.39	15.34	6.78	11.48	11.66	11.61	10.19	12.82	7.85	9.08	11.35	10.82	10.28	12.40
	IDT	RFNet	16.74	25.71	11.71	20.03	14.38	12.65	15.20	18.24	13.18	15.45	11.47	10.06	10.75	12.85	10.25	14.58
		mmFormer	18.14	13.63	21.93	23.05	5.58	17.16	11.69	12.72	8.90	18.40	9.29	4.46	16.71	8.61	6.15	13.09
		M2FTrans	19.22	21.06	18.15	14.14	6.23	13.96	9.64	15.12	9.94	14.21	10.11	7.98	13.80	11.28	8.06	12.86
	IDT (+PASSION)	RFNet	10.01	11.34	10.50	12.22	12.91	9.55	10.94	9.29	9.57	11.34	8.57	8.21	11.24	8.51	7.06	10.08
		mmFormer	10.00	6.33	10.79	10.07	6.20	10.08	8.73	9.95	5.51	9.57	8.46	5.83	9.51	7.41	7.86	8.42
		M2FTrans	9.28	7.28	9.99	8.92	5.68	10.54	7.94	7.72	4.87	10.88	7.62	4.68	10.23	6.91	6.47	7.93